

TITLE OF THE INVENTION

PROTEIN-PROTEIN INTERACTIONS

CROSS-REFERENCE TO RELATED APPLICATIONS

- 5 [0001] The present application is related to U.S. provisional patent application Serial No. 60/256,983, filed on 21 December 2000, incorporated herein by reference, and claims priority thereto under 35 USC §119(e).

BACKGROUND OF THE INVENTION

- 10 [0002] The present invention relates to the discovery of novel protein-protein interactions that are involved in mammalian physiological pathways, including physiological disorders or diseases. Examples of physiological disorders and diseases include non-insulin dependent diabetes mellitus (NIDDM), neurodegenerative disorders, such as Alzheimer's Disease (AD), and the like. Thus, the present invention is directed to complexes of these proteins and/or their fragments, 15 antibodies to the complexes, diagnosis of physiological generative disorders (including diagnosis of a predisposition to and diagnosis of the existence of the disorder), drug screening for agents which modulate the interaction of proteins described herein, and identification of additional proteins in the pathway common to the proteins described herein.

- 20 [0003] The publications and other materials used herein to illuminate the background of the invention, and in particular, cases to provide additional details respecting the practice, are incorporated herein by reference, and for convenience, are referenced by author and date in the following text and respectively grouped in the appended Bibliography.

- 25 [0004] Many processes in biology, including transcription, translation and metabolic or signal transduction pathways, are mediated by non-covalently associated protein complexes. The formation of protein-protein complexes or protein-DNA complexes produce the most efficient chemical machinery. Much of modern biological research is concerned with identifying proteins involved in cellular processes, determining their functions, and how, when and where they interact with other proteins involved in specific pathways. Further, with rapid advances in genome sequencing, there is a need to define protein linkage maps, i.e., detailed inventories of protein 30 interactions that make up functional assemblies of proteins or protein complexes or that make up physiological pathways.

[0005] Recent advances in human genomics research has led to rapid progress in the identification of novel genes. In applications to biological and pharmaceutical research, there is a

need to determine functions of gene products. A first step in defining the function of a novel gene is to determine its interactions with other gene products in appropriate context. That is, since proteins make specific interactions with other proteins or other biopolymers as part of functional assemblies or physiological pathways, an appropriate way to examine function of a gene is to determine its physical relationship with other genes. Several systems exist for identifying protein interactions and hence relationships between genes.

[0006] There continues to be a need in the art for the discovery of additional protein-protein interactions involved in mammalian physiological pathways. There continues to be a need in the art also to identify the protein-protein interactions that are involved in mammalian physiological disorders and diseases, and to thus identify drug targets.

SUMMARY OF THE INVENTION

[0007] The present invention relates to the discovery of protein-protein interactions that are involved in mammalian physiological pathways, including physiological disorders or diseases, and to the use of this discovery. The identification of the interacting proteins described herein provide new targets for the identification of useful pharmaceuticals, new targets for diagnostic tools in the identification of individuals at risk, sequences for production of transformed cell lines, cellular models and animal models, and new bases for therapeutic intervention in such physiological pathways

[0008] Thus, one aspect of the present invention is protein complexes. The protein complexes are a complex of (a) two interacting proteins, (b) a first interacting protein and a fragment of a second interacting protein, (c) a fragment of a first interacting protein and a second interacting protein, or (d) a fragment of a first interacting protein and a fragment of a second interacting protein. The fragments of the interacting proteins include those parts of the proteins, which interact to form a complex. This aspect of the invention includes the detection of protein interactions and the production of proteins by recombinant techniques. The latter embodiment also includes cloned sequences, vectors, transfected or transformed host cells and transgenic animals.

[0009] A second aspect of the present invention is an antibody that is immunoreactive with the above complex. The antibody may be a polyclonal antibody or a monoclonal antibody. While the antibody is immunoreactive with the complex, it is not immunoreactive with the component parts of the complex. That is, the antibody is not immunoreactive with a first interactive protein, a fragment of a first interacting protein, a second interacting protein or a fragment of a second

interacting protein. Such antibodies can be used to detect the presence or absence of the protein complexes.

[0010] A third aspect of the present invention is a method for diagnosing a predisposition for physiological disorders or diseases in a human or other animal. The diagnosis of such disorders includes a diagnosis of a predisposition to the disorders and a diagnosis for the existence of the disorders. In accordance with this method, the ability of a first interacting protein or fragment thereof to form a complex with a second interacting protein or a fragment thereof is assayed, or the genes encoding interacting proteins are screened for mutations in interacting portions of the protein molecules. The inability of a first interacting protein or fragment thereof to form a complex, or the presence of mutations in a gene within the interacting domain, is indicative of a predisposition to, or existence of a disorder. In accordance with one embodiment of the invention, the ability to form a complex is assayed in a two-hybrid assay. In a first aspect of this embodiment, the ability to form a complex is assayed by a yeast two-hybrid assay. In a second aspect, the ability to form a complex is assayed by a mammalian two-hybrid assay. In a second embodiment, the ability to form a complex is assayed by measuring *in vitro* a complex formed by combining said first protein and said second protein. In one aspect the proteins are isolated from a human or other animal. In a third embodiment, the ability to form a complex is assayed by measuring the binding of an antibody, which is specific for the complex. In a fourth embodiment, the ability to form a complex is assayed by measuring the binding of an antibody that is specific for the complex with a tissue extract from a human or other animal. In a fifth embodiment, coding sequences of the interacting proteins described herein are screened for mutations.

[0011] A fourth aspect of the present invention is a method for screening for drug candidates which are capable of modulating the interaction of a first interacting protein and a second interacting protein. In this method, the amount of the complex formed in the presence of a drug is compared with the amount of the complex formed in the absence of the drug. If the amount of complex formed in the presence of the drug is greater than or less than the amount of complex formed in the absence of the drug, the drug is a candidate for modulating the interaction of the first and second interacting proteins. The drug promotes the interaction if the complex formed in the presence of the drug is greater and inhibits (or disrupts) the interaction if the complex formed in the presence of the drug is less. The drug may affect the interaction directly, i.e., by modulating the binding of the two proteins, or indirectly, e.g., by modulating the expression of one or both of the proteins.

[0012] A fifth aspect of the present invention is a model for such physiological pathways, disorders or diseases. The model may be a cellular model or an animal model, as further described herein. In accordance with one embodiment of the invention, an animal model is prepared by creating transgenic or “knock-out” animals. The knock-out may be a total knock-out, i.e., the desired gene is deleted, or a conditional knock-out, i.e., the gene is active until it is knocked out at a determined time. In a second embodiment, a cell line is derived from such animals for use as a model. In a third embodiment, an animal model is prepared in which the biological activity of a protein complex of the present invention has been altered. In one aspect, the biological activity is altered by disrupting the formation of the protein complex, such as by the binding of an antibody or small molecule to one of the proteins which prevents the formation of the protein complex. In a second aspect, the biological activity of a protein complex is altered by disrupting the action of the complex, such as by the binding of an antibody or small molecule to the protein complex which interferes with the action of the protein complex as described herein. In a fourth embodiment, a cell model is prepared by altering the genome of the cells in a cell line. In one aspect, the genome of the cells is modified to produce at least one protein complex described herein. In a second aspect, the genome of the cells is modified to eliminate at least one protein of the protein complexes described herein.

[0013] A sixth aspect of the present invention are nucleic acids coding for novel proteins discovered in accordance with the present invention and the corresponding proteins and antibodies.

[0014] A seventh aspect of the present invention is a method of screening for drug candidates useful for treating a physiological disorder. In this embodiment, drugs are screened on the basis of the association of a protein with a particular physiological disorder. This association is established in accordance with the present invention by identifying a relationship of the protein with a particular physiological disorder. The drugs are screened by comparing the activity of the protein in the presence and absence of the drug. If a difference in activity is found, then the drug is a drug candidate for the physiological disorder. The activity of the protein can be assayed *in vitro* or *in vivo* using conventional techniques, including transgenic animals and cell lines of the present invention.

DETAILED DESCRIPTION OF THE INVENTION

[0015] The present invention is the discovery of novel interactions between proteins described herein. The genes coding for some of these proteins may have been cloned previously,

but their potential interaction in a physiological pathway or with a particular protein was unknown. Alternatively, the genes coding for some of these proteins have not been cloned previously and represent novel genes. These proteins are identified using the yeast two-hybrid method and searching a human total brain library, as more fully described below.

[0016] According to the present invention, new protein-protein interactions have been discovered. The discovery of these interactions has identified several protein complexes for each protein-protein interaction. The protein complexes for these interactions are set forth below in Tables 1-10, which also identifies the new protein-protein interactions of the present invention.

TABLE 1

Protein Complexes LXR-alpha/Utrophin Interaction

Oxysterol liver X receptor alpha (LXR-alpha) and utrophin
 A fragment of LXR-alpha and utrophin
 LXR-alpha and a fragment of utrophin
 A fragment of LXR-alpha and a fragment of utrophin

TABLE 2

Protein Complexes LXR-alpha/zyxin Interaction

Oxysterol liver X receptor alpha (LXR-alpha) and zyxin
 A fragment of LXR-alpha and zyxin
 LXR-alpha and a fragment of zyxin
 A fragment of LXR-alpha and a fragment of zyxin

TABLE 3

Protein Complexes LXR-alpha/LIMS1 Interaction

Oxysterol liver X receptor alpha (LXR-alpha) and LIMS1
 A fragment of LXR-alpha and LIMS1
 LXR-alpha and a fragment of LIMS1
 A fragment of LXR-alpha and a fragment of LIMS1

TABLE 4

Protein Complexes LXR-alpha/PN7771 Interaction

Oxysterol liver X receptor alpha (LXR-alpha) and PN7771

A fragment of LXR-alpha and PN7771

5 LXR-alpha and a fragment of PN7771

A fragment of LXR-alpha and a fragment of PN7771

TABLE 5

Protein Complexes LXR-alpha/Homer-3 Interaction

10 Oxysterol liver X receptor alpha (LXR-alpha) and Homer-3

A fragment of LXR-alpha and Homer-3

LXR-alpha and a fragment of Homer-3

A fragment of LXR-alpha and a fragment of Homer-3

TABLE 6

Protein Complexes LXR-alpha/RACK1 Interaction

Oxysterol liver X receptor alpha (LXR-alpha) and RACK1

A fragment of LXR-alpha and RACK1

LXR-alpha and a fragment of RACK1

20 A fragment of LXR-alpha and a fragment of RACK1

TABLE 7

Protein Complexes LXR-alpha/EIF3S1 Interaction

Oxysterol liver X receptor alpha (LXR-alpha) and EIF3S1

25 A fragment of LXR-alpha and EIF3S1

LXR-alpha and a fragment of EIF3S1

A fragment of LXR-alpha and a fragment of EIF3S1

TABLE 8

Protein Complexes LXR-alpha/PSMD11 Interaction

Oxysterol liver X receptor alpha (LXR-alpha) and PSMD11

A fragment of LXR-alpha and PSMD11

LXR-alpha and a fragment of PSMD11

A fragment of LXR-alpha and a fragment of PSMD11

TABLE 9

Protein Complexes LXR-alpha/KIAA0610 Interaction

Oxysterol liver X receptor alpha (LXR-alpha) and KIAA0610

A fragment of LXR-alpha and KIAA0610

LXR-alpha and a fragment of KIAA0610

A fragment of LXR-alpha and a fragment of KIAA0610

TABLE 10

Protein Complexes LXR-alpha/CIR Interaction

Oxysterol liver X receptor alpha (LXR-alpha) and CIR

A fragment of LXR-alpha and CIR

LXR-alpha and a fragment of CIR

A fragment of LXR-alpha and a fragment of CIR

[0017] The involvement of above interactions in particular pathways is as follows.

[0018] Many cellular proteins exert their function by interacting with other proteins in the cell. Examples of this are found in the formation of multiprotein complexes and the association of enzymes with their substrates. It is widely believed that a great deal of information can be gained by understanding individual protein-protein interactions, and that this is useful in identifying complex networks of interacting proteins that participate in the workings of normal cellular functions. Ultimately, the knowledge gained by characterizing these networks can lead to valuable insight into the causes of human diseases and can eventually lead to the development of therapeutic strategies. The yeast two-hybrid assay is a powerful tool for determining protein-protein interactions and it has been successfully used for studying human disease pathways. In one variation of this technique, a protein of interest (or a portion of that protein) is expressed in a

population of yeast cells that collectively contain all protein sequences. Yeast cells that possess protein sequences that interact with the protein of interest are then genetically selected, and the identity of those interacting proteins are determined by DNA sequencing. Thus, proteins that can be demonstrated to interact with a protein known to be involved in a human disease are therefore also implicated in that disease. Proteins identified in the first round of two-hybrid screening can be subsequently used in a second round of two-hybrid screening, allowing the identification of multiple proteins in the complex network of interactions in a disease pathway.

[0019] Nuclear hormone receptors play important roles in development, reproduction, and physiology by altering gene transcription in response to hormonal signals (Whitfield et al., 1999; Klein-Hitpass et al., 1998). Misregulation of hormone receptor signaling pathways is responsible for a variety of diseases. For example, aldosterone and its receptor (the mineralocorticoid receptor, MCR) are involved in hypertension and congestive heart failure (Duprez et al., 2000), and it has recently been shown that a missense mutation in MCR that alters its ligand specificity is responsible for pregnancy-exacerbated hypertension (Geller et al., 2000). Likewise, glucocorticoids and the glucocorticoid receptor (GR) have been implicated in chronic inflammation and arthritis (Banres, 1998), and the oxysterol liver receptor (LXR), farnesoid X receptor (FXR), and other nuclear receptors are involved in cholesterol homeostasis and atherogenesis (Schroepfer, 2000; Haynes et al., 2000; Brown and Jessup, 1999)

[0020] Collectively, the nuclear receptor superfamily is responsive to a wide variety of ligands. Nuclear hormone receptors share several important structural features, including a variable N-terminal region, a conserved central DNA-binding domain, a variable hinge region, and a conserved C-terminal ligand-binding domain (Moras and Gronemeyer, 1998; Mangelsdorf et al., 1995). Despite this conserved structural organization, interactions between ligands and receptors are remarkably specific. Hormone binding results in conformational changes in the receptor, allowing binding to specific DNA sequences (hormone response elements, HREs) in target gene promoters resulting in changes in target gene transcription. Interaction of nuclear hormone receptors with accessory proteins determines whether the receptor activates or represses transcription. Receptors can recruit coactivators that remodel chromatin and stabilize the RNA polymerase machinery, or alternatively can interact with factors that condense chromatin structure and inactivate gene expression (Wolffe et al., 1997). Furthermore, binding of a nuclear hormone receptor to other cellular proteins can alter the subcellular localization of the receptor and control its ability to bind hormone and HREs (DeFranco et al., 1998). Clearly, identification of factors with which nuclear

hormone receptors interact is vital to understanding the process by which hormonal signals are transduced into transcriptional responses. In addition, identification of receptor-interacting proteins will increase the repertoire of potential targets for therapeutic intervention in the treatment of diseases due to defects involving nuclear hormone signaling.

[0021] The oxysterol liver X receptor alpha (LXR α) was used in yeast two-hybrid searches to identify novel protein-protein interactions. Here, we describe ten new interactors for LXR α . The first four interactors are involved in cell adhesion and cellular architecture. The first of these is the actin-binding protein utrophin. Utrophin is an autosomal gene that is similar to dystrophin, the gene famous for its role in Duchenne's muscular dystrophy. Unlike, dystrophin, however, utrophin appears to be expressed in a wide variety of adult tissues. Dystrophin and the dystrophin-related proteins contain spectrin repeats and likely play a role in anchoring the cytoskeleton to the plasma membrane by their actin-binding activities. The second interactor is the adhesion plaque protein zyxin, also involved in anchoring the cytoskeleton. Zyxin is a phosphoprotein that contains three LIM domains and two proline-rich regions. The interaction of LXR α with zyxin is reminiscent of the interaction we have identified between the farnesoid X-activated receptor and the LIM domain cytoskeletal protein Paxillin. Third, LXR α interacts with the novel protein PN7771, which is highly related (greater than 90% amino acid identity) to Ninein. Ninein is a centrosome-associated protein that interacts with human glycogen synthase kinase 3 β (GSK-3 β) (Hong et al., 2000), is localized to the pericentriolar matrix of the centrosome, and reacts with centrosomal autoantibody sera (Mack et al., 1998). PN7771 contains predicted calcium-binding EF hand motifs, a potential nuclear localization signal, a basic region-leucine zipper motif, a spectrin repeat, coiled-coil motifs, and Glu- and Gln-rich regions. Taken together, these interactions suggest that LXR α may be involved in cellular signaling events in response to cellular adhesion or other extracellular stimuli, and that the trans-activating ability of LXR α may be regulated by its interaction with these proteins.

[0022] Several LXR α interactors are involved in signal transduction pathways. The first is the neuronal immediate early protein homer-3. Homer proteins bind to the C-terminal tails of metabotropic glutamate receptors and play a role in their targeting and regulation; the metabotropic glutamate receptors, in turn, participate in the influx of intracellular calcium. Since LXR- α binds to homer-3, it is possible that LXR α may also be involved in calcium release. Alternatively, LXR- α could be modulated in some way by homer-3 in a manner analogous to the way in which the metabotropic glutamate receptors are regulated. The second protein, RACK1 (receptor of activated protein kinase C 1), is a WD repeat-containing protein that functions as an intracellular

receptor to localize PKC to the cytoskeleton. The interaction between RACK1 and LXR-alpha suggests that LXR-alpha may be capable of localizing to the cytoskeleton via its association with RACK1. The next interactor is the LIM-domain protein LIMS1. LIMS1 has been implicated in integrin-linked kinase signaling, and it has been shown to interact with the SH3 and SH2 domain-containing adaptor protein NCK2 (Tu et al. 1998). Taken together, these findings suggest the involvement of LXRa in a variety of signal transduction pathways; whether LXRa activity is regulated by interaction with these proteins, or vice versa, remains to be determined.

[0023] Two LXRa interactors involved in protein metabolism were identified: the proteasome subunit PSMD11, which is involved in protein turnover, and the translation initiation factor EIF3S1, which is involved in protein synthesis. The interaction of these proteins with LXRa suggests that nuclear hormone receptors may be involved directly with protein production and stability, in addition to transcriptional regulation.

[0024] An interaction between LXRa and the potential transmembrane protein KIAA0610 was identified. KIAA0610 is hypothetical protein fragment 686 amino acids in length. Predicted structural motifs include four possible transmembrane domains and a coiled-coil domain. KIAA0610 displays weak homology (~24% amino acid identity over 360-430 residues) to *Drosophila* and *C. elegans* proteins of unknown function. EST analysis suggests expression in a variety of tissues.

[0025] Finally, an interaction between LXRa and the transcription factor CIR was identified. CIR has been demonstrated to interact with the CBF1 transcription factor as well as histone deacetylase HD2 and Sin3-associated protein 30kD (Hsieh et al., 1999). It has been proposed that CIR acts as a linker between CBF1 and the histone deacetylase complex. Similarly, the interaction of LXRa with CIR suggests CIR may link LXRa with the histone deacetylase machinery. In support of a functional role between nuclear receptors and CIR, we have also identified an interaction between CIR and the estrogen receptor ER-beta.

[0026] The proteins disclosed in the present invention were found to interact with their corresponding proteins in the yeast two-hybrid system. Because of the involvement of the corresponding proteins in the physiological pathways disclosed herein, the proteins disclosed herein also participate in the same physiological pathways. Therefore, the present invention provides a list of uses of these proteins and DNA encoding these proteins for the development of diagnostic and therapeutic tools useful in the physiological pathways. This list includes, but is not limited to, the following examples.

Two-Hybrid System

[0027] The principles and methods of the yeast two-hybrid system have been described in detail elsewhere (e.g., Bartel and Fields, 1997; Bartel et al., 1993; Fields and Song, 1989; Chevray and Nathans, 1992). The following is a description of the use of this system to identify proteins that interact with a protein of interest.

[0028] The target protein is expressed in yeast as a fusion to the DNA-binding domain of the yeast Gal4p. DNA encoding the target protein or a fragment of this protein is amplified from cDNA by PCR or prepared from an available clone. The resulting DNA fragment is cloned by ligation or recombination into a DNA-binding domain vector (e.g., pGBT9, pGBT.C, pAS2-1) such that an in-frame fusion between the Gal4p and target protein sequences is created.

[0029] The target gene construct is introduced, by transformation, into a haploid yeast strain. A library of activation domain fusions (i.e., adult brain cDNA cloned into an activation domain vector) is introduced by transformation into a haploid yeast strain of the opposite mating type. The yeast strain that carries the activation domain constructs contains one or more Gal4p-responsive reporter gene(s), whose expression can be monitored. Examples of some yeast reporter strains include Y190, PJ69, and CBY14a. An aliquot of yeast carrying the target gene construct is combined with an aliquot of yeast carrying the activation domain library. The two yeast strains mate to form diploid yeast and are plated on media that selects for expression of one or more Gal4p-responsive reporter genes. Colonies that arise after incubation are selected for further characterization.

[0030] The activation domain plasmid is isolated from each colony obtained in the two-hybrid search. The sequence of the insert in this construct is obtained by the dideoxy nucleotide chain termination method. Sequence information is used to identify the gene/protein encoded by the activation domain insert via analysis of the public nucleotide and protein databases. Interaction of the activation domain fusion with the target protein is confirmed by testing for the specificity of the interaction. The activation domain construct is co-transformed into a yeast reporter strain with either the original target protein construct or a variety of other DNA-binding domain constructs. Expression of the reporter genes in the presence of the target protein but not with other test proteins indicates that the interaction is genuine.

[0031] In addition to the yeast two-hybrid system, other genetic methodologies are available for the discovery or detection of protein-protein interactions. For example, a mammalian two-hybrid

system is available commercially (Clontech, Inc.) that operates on the same principle as the yeast two-hybrid system. Instead of transforming a yeast reporter strain, plasmids encoding DNA-binding and activation domain fusions are transfected along with an appropriate reporter gene (e.g., *lacZ*) into a mammalian tissue culture cell line. Because transcription factors such as the *Saccharomyces cerevisiae* Gal4p are functional in a variety of different eukaryotic cell types, it would be expected that a two-hybrid assay could be performed in virtually any cell line of eukaryotic origin (e.g., insect cells (SF9), fungal cells, worm cells, etc.). Other genetic systems for the detection of protein-protein interactions include the so-called SOS recruitment system (Aronheim et al., 1997).

10 Protein-protein interactions

[0032] Protein interactions are detected in various systems including the yeast two-hybrid system, affinity chromatography, co-immunoprecipitation, subcellular fractionation and isolation of large molecular complexes. Each of these methods is well characterized and can be readily performed by one skilled in the art. See, e.g., U.S. Patents No. 5,622,852 and 5,773,218, and PCT published applications No. WO 97/27296 and WO 99/65939, each of which are incorporated herein by reference.

[0033] The protein of interest can be produced in eukaryotic or prokaryotic systems. A cDNA encoding the desired protein is introduced in an appropriate expression vector and transfected in a host cell (which could be bacteria, yeast cells, insect cells, or mammalian cells). Purification of the expressed protein is achieved by conventional biochemical and immunochemical methods well known to those skilled in the art. The purified protein is then used for affinity chromatography studies: it is immobilized on a matrix and loaded on a column. Extracts from cultured cells or homogenized tissue samples are then loaded on the column in appropriate buffer, and non-binding proteins are eluted. After extensive washing, binding proteins or protein complexes are eluted using various methods such as a gradient of pH or a gradient of salt concentration. Eluted proteins can then be separated by two-dimensional gel electrophoresis, eluted from the gel, and identified by micro-sequencing. The purified proteins can also be used for affinity chromatography to purify interacting proteins disclosed herein. All of these methods are well known to those skilled in the art.

[0034] Similarly, both proteins of the complex of interest (or interacting domains thereof) can be produced in eukaryotic or prokaryotic systems. The proteins (or interacting domains) can be under control of separate promoters or can be produced as a fusion protein. The fusion protein

may include a peptide linker between the proteins (or interacting domains) which, in one embodiment, serves to promote the interaction of the proteins (or interacting domains). All of these methods are also well known to those skilled in the art.

[0035] Purified proteins of interest, individually or a complex, can also be used to generate antibodies in rabbit, mouse, rat, chicken, goat, sheep, pig, guinea pig, bovine, and horse. The methods used for antibody generation and characterization are well known to those skilled in the art. Monoclonal antibodies are also generated by conventional techniques. Single chain antibodies are further produced by conventional techniques.

[0036] DNA molecules encoding proteins of interest can be inserted in the appropriate expression vector and used for transfection of eukaryotic cells such as bacteria, yeast, insect cells, or mammalian cells, following methods well known to those skilled in the art. Transfected cells expressing both proteins of interest are then lysed in appropriate conditions, one of the two proteins is immunoprecipitated using a specific antibody, and analyzed by polyacrylamide gel electrophoresis. The presence of the binding protein (co-immunoprecipitated) is detected by immunoblotting using an antibody directed against the other protein. Co-immunoprecipitation is a method well known to those skilled in the art.

[0037] Transfected eukaryotic cells or biological tissue samples can be homogenized and fractionated in appropriate conditions that will separate the different cellular components. Typically, cell lysates are run on sucrose gradients, or other materials that will separate cellular components based on size and density. Subcellular fractions are analyzed for the presence of proteins of interest with appropriate antibodies, using immunoblotting or immunoprecipitation methods. These methods are all well known to those skilled in the art.

Disruption of protein-protein interactions

[0038] It is conceivable that agents that disrupt protein-protein interactions can be beneficial in many physiological disorders, including, but not-limited to NIDDM, AD and others disclosed herein. Each of the methods described above for the detection of a positive protein-protein interaction can also be used to identify drugs that will disrupt said interaction. As an example, cells transfected with DNAs coding for proteins of interest can be treated with various drugs, and co-immunoprecipitations can be performed. Alternatively, a derivative of the yeast two-hybrid system, called the reverse yeast two-hybrid system (Leanna and Hannink, 1996), can be used, provided that the two proteins interact in the straight yeast two-hybrid system.

Modulation of protein-protein interactions

[0039] Since the interactions described herein are involved in a physiological pathway, the identification of agents which are capable of modulating the interactions will provide agents which can be used to track physiological disorder or to use lead compounds for development of therapeutic agents. An agent may modulate expression of the genes of interacting proteins, thus affecting interaction of the proteins. Alternatively, the agent may modulate the interaction of the proteins.

The agent may modulate the interaction of wild-type with wild-type proteins, wild-type with mutant proteins, or mutant with mutant proteins. Agents which may be used to modulate the protein interaction include a peptide, an antibody, a nucleic acid, an antisense compound or a ribozyme.

The nucleic acid may encode the antibody or the antisense compound. The peptide may be at least 4 amino acids of the sequence of either of the interacting proteins. Alternatively, the peptide may be from 4 to 30 amino acids (or from 8 to 20 amino acids) that is at least 75% identical to a contiguous span of amino acids of either of the interacting proteins. The peptide may be covalently linked to a transporter capable of increasing cellular uptake of the peptide. Examples of a suitable transporter include penetratins, *l*-Tat₄₉₋₅₇, *d*-Tat₄₉₋₅₇, retro-inverso isomers of *l*- or *d*-Tat₄₉₋₅₇, L-arginine oligomers, D-arginine oligomers, L-lysine oligomers, D-lysine oligomers, L-histidine oligomers, D-histidine oligomers, L-ornithine oligomers, D-ornithine oligomers, short peptide sequences derived from fibroblast growth factor, Galparan, and HSV-1 structural protein VP22, and peptoid analogs thereof. Agents can be tested using transfected host cells, cell lines, cell models or animals, such as described herein, by techniques well known to those of ordinary skill in the art, such as disclosed in U.S. Patents Nos. 5,622,852 and 5,773,218, and PCT published application Nos. WO 97/27296 and WO 99/65939, each of which are incorporated herein by reference. The modulating effect of the agent can be tested *in vivo* or *in vitro*. Agents can be provided for testing in a phage display library or a combinatorial library. Exemplary of a method to screen agents is to measure the effect that the agent has on the formation of the protein complex.

Mutation screening

[0040] The proteins disclosed in the present invention interact with one or more proteins known to be involved in a physiological pathway, such as in NIDDM, AD or pathways described herein. Mutations in interacting proteins could also be involved in the development of the physiological disorder, such as NIDDM, AD or disorders described herein, for example, through

a modification of protein-protein interaction, or a modification of enzymatic activity, modification of receptor activity, or through an unknown mechanism. Therefore, mutations can be found by sequencing the genes for the proteins of interest in patients having the physiological disorder, such as insulin, and non-affected controls. A mutation in these genes, especially in that portion of the gene involved in protein interactions in the physiological pathway, can be used as a diagnostic tool and the mechanistic understanding the mutation provides can help develop a therapeutic tool.

Screening for at-risk individuals

[0041] Individuals can be screened to identify those at risk by screening for mutations in the protein disclosed herein and identified as described above. Alternatively, individuals can be screened by analyzing the ability of the proteins of said individual disclosed herein to form natural complexes. Further, individuals can be screened by analyzing the levels of the complexes or individual proteins of the complexes or the mRNA encoding the protein members of the complexes. Techniques to detect the formation of complexes, including those described above, are known to those skilled in the art. Techniques and methods to detect mutations are well known to those skilled in the art. Techniques to detect the level of the complexes, proteins or mRNA are well known to those skilled in the art.

Cellular models of Physiological Disorders

[0042] A number of cellular models of many physiological disorders or diseases have been generated. The presence and the use of these models are familiar to those skilled in the art. As an example, primary cell cultures or established cell lines can be transfected with expression vectors encoding the proteins of interest, either wild-type proteins or mutant proteins. The effect of the proteins disclosed herein on parameters relevant to their particular physiological disorder or disease can be readily measured. Furthermore, these cellular systems can be used to screen drugs that will influence those parameters, and thus be potential therapeutic tools for the particular physiological disorder or disease. Alternatively, instead of transfecting the DNA encoding the protein of interest, the purified protein of interest can be added to the culture medium of the cells under examination, and the relevant parameters measured.

Animal models

[0043] The DNA encoding the protein of interest can be used to create animals that overexpress said protein, with wild-type or mutant sequences (such animals are referred to as “transgenic”), or animals which do not express the native gene but express the gene of a second animal (referred to as “transplacement”), or animals that do not express said protein (referred to as “knock-out”). The knock-out animal may be an animal in which the gene is knocked out at a determined time. The generation of transgenic, transplacement and knock-out animals (normal and conditioned) uses methods well known to those skilled in the art.

[0044] In these animals, parameters relevant to the particular physiological disorder can be measured. These parameters may include receptor function, protein secretion *in vivo* or *in vitro*, survival rate of cultured cells, concentration of particular protein in tissue homogenates, signal transduction, behavioral analysis, protein synthesis, cell cycle regulation, transport of compounds across cell or nuclear membranes, enzyme activity, oxidative stress, production of pathological products, and the like. The measurements of biochemical and pathological parameters, and of behavioral parameters, where appropriate, are performed using methods well known to those skilled in the art. These transgenic, transplacement and knock-out animals can also be used to screen drugs that may influence the biochemical, pathological, and behavioral parameters relevant to the particular physiological disorder being studied. Cell lines can also be derived from these animals for use as cellular models of the physiological disorder, or in drug screening.

Rational drug design

[0045] The goal of rational drug design is to produce structural analogs of biologically active polypeptides of interest or of small molecules with which they interact (e.g., agonists, antagonists, inhibitors) in order to fashion drugs which are, for example, more active or stable forms of the polypeptide, or which, e.g., enhance or interfere with the function of a polypeptide *in vivo*. Several approaches for use in rational drug design include analysis of three-dimensional structure, alanine scans, molecular modeling and use of anti-id antibodies. These techniques are well known to those skilled in the art. Such techniques may include providing atomic coordinates defining a three-dimensional structure of a protein complex formed by said first polypeptide and said second polypeptide, and designing or selecting compounds capable of interfering with the interaction between a first polypeptide and a second polypeptide based on said atomic coordinates.

[0046] Following identification of a substance which modulates or affects polypeptide activity, the substance may be further investigated. Furthermore, it may be manufactured and/or used in preparation, i.e., manufacture or formulation, or a composition such as a medicament, pharmaceutical composition or drug. These may be administered to individuals.

5 [0047] A substance identified as a modulator of polypeptide function may be peptide or non-peptide in nature. Non-peptide “small molecules” are often preferred for many *in vivo* pharmaceutical uses. Accordingly, a mimetic or mimic of the substance (particularly if a peptide) may be designed for pharmaceutical use.

10 [0048] The designing of mimetics to a known pharmaceutically active compound is a known approach to the development of pharmaceuticals based on a “lead” compound. This approach might be desirable where the active compound is difficult or expensive to synthesize or where it is unsuitable for a particular method of administration, e.g., pure peptides are unsuitable active agents for oral compositions as they tend to be quickly degraded by proteases in the alimentary canal. Mimetic design, synthesis and testing is generally used to avoid randomly screening large numbers
15 of molecules for a target property.

[0049] Once the pharmacophore has been found, its structure is modeled according to its physical properties, e.g., stereochemistry, bonding, size and/or charge, using data from a range of sources, e.g., spectroscopic techniques, x-ray diffraction data and NMR. Computational analysis, similarity mapping (which models the charge and/or volume of a pharmacophore, rather than the
20 bonding between atoms) and other techniques can be used in this modeling process.

[0050] A template molecule is then selected, onto which chemical groups that mimic the pharmacophore can be grafted. The template molecule and the chemical groups grafted thereon can be conveniently selected so that the mimetic is easy to synthesize, is likely to be pharmacologically acceptable, and does not degrade *in vivo*, while retaining the biological activity of the lead
25 compound. Alternatively, where the mimetic is peptide-based, further stability can be achieved by cyclizing the peptide, increasing its rigidity. The mimetic or mimetics found by this approach can then be screened to see whether they have the target property, or to what extent it is exhibited. Further optimization or modification can then be carried out to arrive at one or more final mimetics for *in vivo* or clinical testing.

Diagnostic Assays

[0051] The identification of the interactions disclosed herein enables the development of diagnostic assays and kits, which can be used to determine a predisposition to or the existence of a physiological disorder. In one aspect, one of the proteins of the interaction is used to detect the presence of a “normal” second protein (i.e., normal with respect to its ability to interact with the first protein) in a cell extract or a biological fluid, and further, if desired, to detect the quantitative level of the second protein in the extract or biological fluid. The absence of the “normal” second protein would be indicative of a predisposition or existence of the physiological disorder. In a second aspect, an antibody against the protein complex is used to detect the presence and/or quantitative level of the protein complex. The absence of the protein complex would be indicative of a predisposition or existence of the physiological disorder.

Nucleic Acids and Proteins

[0052] A nucleic acid or fragment thereof has substantial identity with another if, when optimally aligned (with appropriate nucleotide insertions or deletions) with the other nucleic acid (or its complementary strand), there is nucleotide sequence identity in at least about 60% of the nucleotide bases, usually at least about 70%, more usually at least about 80%, preferably at least about 90%, more preferably at least about 95% of the nucleotide bases, and more preferably at least about 98% of the nucleotide bases. A protein or fragment thereof has substantial identity with another if, optimally aligned, there is an amino acid sequence identity of at least about 30% identity with an entire naturally-occurring protein or a portion thereof, usually at least about 70% identity, more usually at least about 80% identity, preferably at least about 90% identity, more preferably at least about 95% identity, and most preferably at least about 98% identity.

[0053] Identity means the degree of sequence relatedness between two polypeptide or two polynucleotides sequences as determined by the identity of the match between two strings of such sequences. Identity can be readily calculated. While there exist a number of methods to measure identity between two polynucleotide or polypeptide sequences, the term "identity" is well known to skilled artisans (*Computational Molecular Biology*, Lesk, A. M., ed., Oxford University Press, New York, 1988; *Biocomputing: Informatics and Genome Projects*, Smith, D. W., ed., Academic Press, New York, 1993; *Computer Analysis of Sequence Data*, Part I, Griffin, A. M., and Griffin, H. G., eds., Humana Press, New Jersey, 1994; *Sequence Analysis in Molecular Biology*, von Heinje, G., Academic Press, 1987; and *Sequence Analysis Primer*, Gribskov, M. and Devereux, J., eds., M

Stockton Press, New York, 1991). Methods commonly employed to determine identity between two sequences include, but are not limited to those disclosed in *Guide to Huge Computers*, Martin J. Bishop, ed., Academic Press, San Diego, 1994, and Carillo, H., and Lipman, D., *SIAM J Applied Math.* **48**:1073 (1988). Preferred methods to determine identity are designed to give the largest match between the two sequences tested. Such methods are codified in computer programs. Preferred computer program methods to determine identity between two sequences include, but are not limited to, GCG (Genetics Computer Group, Madison Wis.) program package (Devereux, J., et al., *Nucleic Acids Research* **12**(1).387 (1984)), BLASTP, BLASTN, FASTA (Altschul et al. (1990); Altschul et al. (1997)). The well-known Smith Waterman algorithm may also be used to determine identity.

[0054] Alternatively, substantial homology or similarity exists when a nucleic acid or fragment thereof will hybridize to another nucleic acid (or a complementary strand thereof) under selective hybridization conditions, to a strand, or to its complement. Selectivity of hybridization exists when hybridization which is substantially more selective than total lack of specificity occurs. Nucleic acid hybridization will be affected by such conditions as salt concentration, temperature, or organic solvents, in addition to the base composition, length of the complementary strands, and the number of nucleotide base mismatches between the hybridizing nucleic acids, as will be readily appreciated by those skilled in the art. Stringent temperature conditions will generally include temperatures in excess of 30°C, typically in excess of 37°C, and preferably in excess of 45°C. Stringent salt conditions will ordinarily be less than 1000 mM, typically less than 500 mM, and preferably less than 200 mM. However, the combination of parameters is much more important than the measure of any single parameter. See, e.g., Asubel, 1992; Wetmur and Davidson, 1968.

[0055] The terms "isolated", "substantially pure", and "substantially homogeneous" are used interchangeably to describe a protein or polypeptide which has been separated from components which accompany it in its natural state. A monomeric protein is substantially pure when at least about 60 to 75% of a sample exhibits a single polypeptide sequence. A substantially pure protein will typically comprise about 60 to 90% W/W of a protein sample, more usually about 95%, and preferably will be over about 99% pure. Protein purity or homogeneity may be indicated by a number of means well known in the art, such as polyacrylamide gel electrophoresis of a protein sample, followed by visualizing a single polypeptide band upon staining the gel. For certain

purposes, higher resolution may be provided by using HPLC or other means well known in the art which are utilized for purification.

[0056] Large amounts of the nucleic acids of the present invention may be produced by (a) replication in a suitable host or transgenic animals or (b) chemical synthesis using techniques well known in the art. Constructs prepared for introduction into a prokaryotic or eukaryotic host may comprise a replication system recognized by the host, including the intended polynucleotide fragment encoding the desired polypeptide, and will preferably also include transcription and translational initiation regulatory sequences operably linked to the polypeptide encoding segment. Expression vectors may include, for example, an origin of replication or autonomously replicating sequence (ARS) and expression control sequences, a promoter, an enhancer and necessary processing information sites, such as ribosome-binding sites, RNA splice sites, polyadenylation sites, transcriptional terminator sequences, and mRNA stabilizing sequences. Secretion signals may also be included where appropriate which allow the protein to cross and/or lodge in cell membranes, and thus attain its functional topology, or be secreted from the cell. Such vectors may be prepared by means of standard recombinant techniques well known in the art.

[0057] The nucleic acid or protein may also be incorporated on a microarray. The preparation and use of microarrays are well known in the art. Generally, the microarray may contain the entire nucleic acid or protein, or it may contain one or more fragments of the nucleic acid or protein. Suitable nucleic acid fragments may include at least 17 nucleotides, at least 21 nucleotides, at least 30 nucleotides or at least 50 nucleotides of the nucleic acid sequence, particularly the coding sequence. Suitable protein fragments may include at least 4 amino acids, at least 8 amino acids, at least 12 amino acids, at least 15 amino acids, at least 17 amino acids or at least 20 amino acids. Thus, the present invention is also directed to such nucleic acid and protein fragments.

EXAMPLES

[0058] The present invention is further detailed in the following Examples, which are offered by way of illustration and are not intended to limit the invention in any manner. Standard techniques well known in the art or the techniques specifically described below are utilized.

EXAMPLE 1

Yeast Two-Hybrid System

[0059] The principles and methods of the yeast two-hybrid systems have been described in detail (Bartel and Fields, 1997). The following is thus a description of the particular procedure that we used, which was applied to all proteins.

[0060] The cDNA encoding the bait protein was generated by PCR from brain cDNA. Gene-specific primers were synthesized with appropriate tails added at their 5' ends to allow recombination into the vector pGBTQ. The tail for the forward primer was 5'-GCAGGAAACAGCTATGACCATACAGTCAGCGGCCGCCACC-3' (SEQ ID NO:1) and the tail for the reverse primer was 5'-ACGGCCAGTCGCGTGGAGTGTATGTCATGCGGCCGCTA-3' (SEQ ID NO:2). The tailed PCR product was then introduced by recombination into the yeast expression vector pGBTQ, which is a close derivative of pGBTC (Bartel et al., 1996) in which the polylinker site has been modified to include M13 sequencing sites. The new construct was selected directly in the yeast J693 for its ability to drive tryptophane synthesis (genotype of this strain: Mat α , ade2, his3, leu2, trp1, URA3::GAL1-lacZ LYS2::GAL1-HIS3 gal4del gal80del cyhR2). In these yeast cells, the bait is produced as a C-terminal fusion protein with the DNA binding domain of the transcription factor Gal4 (amino acids 1 to 147). A total human brain (37 year-old male Caucasian) cDNA library cloned into the yeast expression vector pACT2 was purchased from Clontech (human brain MATCHMAKER cDNA, cat. # HL4004AH), transformed into the yeast strain J692 (genotype of this strain: Mat a, ade2, his3, leu2, trp1, URA3::GAL1-lacZ LYS2::GAL1-HIS3 gal4del gal80del cyhR2), and selected for the ability to drive leucine synthesis. In these yeast cells, each cDNA is expressed as a fusion protein with the transcription activation domain of the transcription factor Gal4 (amino acids 768 to 881) and a 9 amino acid hemagglutinin epitope tag. J693 cells (Mat α type) expressing the bait were then mated with J692 cells (Mat a type) expressing proteins from the brain library. The resulting diploid yeast cells expressing proteins interacting with the bait protein were selected for the ability to synthesize tryptophan, leucine, histidine, and β -galactosidase. DNA was prepared from each clone, transformed by electroporation into *E. coli* strain KC8 (Clontech KC8 electrocompetent cells, cat. # C2023-1), and the cells were selected on ampicillin-containing plates in the absence of either tryptophane (selection for the bait plasmid) or leucine (selection for the brain library plasmid). DNA for both plasmids was prepared and sequenced by dideoxynucleotide chain termination method. The identity of the bait cDNA insert was confirmed and the cDNA insert from the brain library plasmid was identified using BLAST program against public

nucleotides and protein databases. Plasmids from the brain library (preys) were then individually transformed into yeast cells together with a plasmid driving the synthesis of lamin fused to the Gal4 DNA binding domain. Clones that gave a positive signal after β -galactosidase assay were considered false-positives and discarded. Plasmids for the remaining clones were transformed into yeast cells together with plasmid for the original bait. Clones that gave a positive signal after β -galactosidase assay were considered true positives.

EXAMPLE 2

Identification of LXR-alpha/Utrophin Interaction

[0061] A yeast two-hybrid system as described in Example 1 using amino acids 95-277 of LXR-alpha (GenBank (GB) accession no. U22662) as bait was performed. One clone that was identified by this procedure included amino acids 2443-2650 of utrophin (GB accession no. X15488).

EXAMPLE 3

Identification of LXR-alpha/PN7771 Interaction

[0062] A yeast two-hybrid system as described in Example 1 using amino acids 95-277 of LXR-alpha (GB accession no. U22662) as bait was performed. One clone that was identified by this procedure included amino acids 1747-2047 of PN7771. The DNA sequence and the predicted protein sequence for PN7771 are set forth in Tables 11 and 12, respectively.

TABLE 11

Nucleotide Sequence of PN7771

```

cttattttgaaaacatttacatagtgattagttaacccaacagaccaatcctgggaagacagccagagcctgcagcaccttagtaacagaaaaactg
ataattaggagaagagacctgtccaagaccaggaacctggacaaaaattgtccatgttgctttactttaatgagtggtgccccagtaaaaactgagct
gtatggcagagctgttcacatttatcttctgtgtccaccagttctgtgaaacccctggcaagatcgtggccctgtgtagcttgcattgtttgaaca
gctgtctatggaagaaagcaaacacaacctagagcaacattgattgttttagaaagctctttattttcagttctggctgtgttcaacatcttagctta
cgttttcatgttgtaatgatctgccgtatggacgatcacctctaagttagagagttctgtaatttggcttggattaaagatcgttggttagtgaagctg
ctgcttttttatagtaaaaggactggttctgagagccttgtgcagatggctgaggtcacctgcccaagggtgatgtcgtgttggcattcattgcat
catggcgaaggcatcttcagatgtgcaggttcaggcttcatcggaatccagcacgttaaaaatgaactttgccatgttgagcttgaggag
gtggccccagtgctgcagcagacattacttcaggacaacctcttgggcagggtacattttgaccaattaaagaagcattaatactcatctgtccag
aactctgtcaaatgaagaacacattcaagaaccagactgctcactagaagctcagcccaatatgttagaggtgggaagcgttacggacgaaggt
ccttggccgagttccaagagtcggtggaggagttcctgaagtgcaggtgattgagccactggatgaagaagcgcggccttcacacatcccagc
cggtgactgcagtgagcactggaagacgcaacgcagtgaggagatgaagcgggaaggccagttaaggttttgaaccagatgacttgaatgc
ttcacagagtggatcttccctcccaagactggatagaagagaaactgcaagaagtttgtgaagatttggggatcaccctgatgtgcacctgaa
ccggaagaagctggtctccatctgtgagcagtatggtttacagaatgtggatggagagatgctcgagggaagtattccataatcttgatcctgacggt
acaatgagtgtagaagatttttctatgtttgttfaaaatggaaaatctctacaccatcagcatctactccatatagacaactaaaaggcaccttc

```

catgcagctttc gatgagagtggacgacgtaccacaacctcatcagcaatgacaagtaccattggctttcgggtcttctcctgcctggatgatggg
 atgggccatgc atctgtggagagaacttgacacctggcaggaagagggcattgagaacagccaggagatcctgaaggccttggatttcagc
 ctcatggaaacatcaatttgacagaaltaacactggcccttgaaaaatgaacttttggttaccaagaacagcattcaccaggcggctctggccagct
 ttaaggctgaaatccggcatttgggtgaacgagttgatcaggtggctcagagaaaaagagaagctacggctcagatctggacaaggccgagaagct
 5 caagtctttaatggcctcggaggtggatgacccatgcggccatagagcggcggaatgagtacaacctcaggaaactggatggagagtacaa
 ggagcgaatagcagccttaaaaaatgaactccgaaaagagagagagcagatcctgcagcaggcaggcaagcagcgtttagaacttgaacagg
 aaattgaaaaggcaaaaacagaagagaactataccgggaccgccttgccctctctttaaggaaaaacagtcgtctggaaaatgagcttctagaa
 aatgcagagaagttggcagaatatgagaatctgacaaacaaactcagagaaattggaaaaatgtgttagcagaaaaagttgggtgacctcgtatcct
 agcagtgtctgagttctcctgcaagaagagagactgacacagatgagaaatgaatatgagcggcagtgagggtactacaagaccaagtagatg
 10 aactccagtctgagctggaagaatatcgtgcacaaggcagagtgctcaggcttccgttgaaagcaccgtcagaagaagttgaggctaacag
 cgggtggcattgagcccgaacacgggctcggttctgaagaatgcaatccattgaatatgagcattgaggcagagctggctcattgaacagatgaaag
 aacaacatcacaggacatgattgacctcagactggagctcgaagataaagtgcgccattatgaaaagcagctggacgaaaccgtggtcagctg
 caagaaggcacaggagaacatgaagcaaaaggcatgagaacgaaacgcgcacctagaaaaacaaataagtaccttaaaaaatgaaattgtga
 acttcaggggcaagcagcagtgctcaaggaggcacatcatgaggccacttcgagggcatgaggaggagaaaaaacaactgcaagtgaagcttg
 15 agggagaaaaagactcacctgcaggagaagctgaggctgcaacatgagatggagctcaaggctagactgacacaggctcaagcaagctttgag
 cgggagagggaaggccttcagagtagcgctggacagaagagaaggtgagaggcttgactcaggaactagagcagttaccaggagcagc
 tgacaagcctgtgtggagaaacacactcttgagaagaggagttgaagaaagagctcttgaaaaagcaccaaaggagcttcaggagggaagg
 gaaaaatgaaacagagtgtatagaagaacctctcaatagaagcccagtttcagctctgattgtcagaagtcactgagaggtgtgaaagcgc
 tctgcaaaagcctggaggggcgctaccgccaaagagctgaaggacctccaggaaacagcagcgtgaggagaaatcccagtgggaaattgagaag
 20 gacgagctcaccaggagtgctgcggaagcccaggagctgctgaaagagactcttaagagagagaaaaacaacttctctggtctgacctaggga
 gagagagatgtgtggagaaacatacaaaagaacatttgaacagcatggtcgtcagagacagcagctactccaagacctggaagacctaaagaaa
 tgtatctgaaaccagcaaaagcctgtgtctgaccagatacttgagctgaagagcagtcacaaaagggaactgagggaagcgtgaggaggtcct
 gtgccaggcaggggcttcggagcagctggccagccagcggctggaaaagactagaaatggaacatgaccaggaaaggcaggaaatgatgtcc
 aagcttctagccatggagaacattcacaagcgacctgtgagacagcagatcgagaagagccgagatgagcacagaaatctccagacttcag
 25 agtaaaaaaaggaaatgcagcaggcaacatctctctcaatgcttcagagtggttgccagggtgataggagaggagaggtggaaggagatg
 gagccctgtccctgcttcagcaaggggagcagctgttggaagaaaatggggacgtcctttaaagcctgcagagagctcatgaacaggcagtg
 aggaaaatgtgaaaatggctactgaaattctagattgcaacagaggctacaaaagttagagccagggttagtaattgtcttctgtttggatgagcca
 gctactgagtttttggaaatactgcggaacaaacagagcagttttacagcaaaaccgaacgaagcaagtagaaggtgtgaccaggcggcatgt
 cctaagtacctggagatgatgaggtccgggacctgggaagtacaggggacgagctctgttcagagacagggaagtcaaaatagaggagctga
 30 agcttcagtagagggttttctgagcttgaacacagtgaagagaccaggactgaatcctgggagctgaagaatcagattagctcagcttcaggaaca
 gctaattgatgtatgtgctggactgtgacgagcttctgaaaaaagaacaggacctactttttagtgttctgtgctaaaaaagaaactgaagatgctga
 gagaatccctgaggcttctccaaatataagctgtgtatgaagatgtgagccgagaaaaatgactgccttcagggaagagctgagaatgatggaga
 cacgctacgatgaggcactagaaaaatacaaaagaactcactgcagaggttttcaggttcaggatgagctgaagaaaatggaggaaagtcactga
 aacattcctcagcctggaaaagagttacgatgaggtcaaaatagaaaatgaggggctgaatgttctggtttgagacttcaaggcaagattgagaa
 35 gcttcaggaaaagcgtggtccagcgggtgtgactgctgtatgggaagccagtttagagaacctggaaatcgaacctgatggaaatatactccagc
 tcaatcagacactggaagagtgtgtcccagggttaggaggtgtacatcatgtcatagagggaatgtaagcaagaaaaccagtaccttgagggggaa
 cacacagctcttgaaaaaagttaaagcacatgaaattgcctgggtacatggaacaattcagacacatcaagaaaggccaaggtacagaatcaa
 gttatactggaggaaaacactactctcctaggcttcaagacaacatttccagc atcaggccaccatagcagagttagaactggagaaaacaaag
 ttacaggagctgactaggaagttgaaggagagagtcactattttagttaagcaaaaagatgtacttttcacggagaaaaaggaggaagagctgaa
 40 ggcaatgatgcatgacttgcatgacacgtgcagtgagatgcagcaaaaagttgaacttctgagatatgaatctgaaaagcttcaacaggaaaattc
 tattttgaaaatgaaattactctttaaataagaagatagcatttcaacctgaaattaggggacattaaatggatctcagggaagaaatgtggcaaaa
 aacggaaactgtaaaaacagaaaatgctgcagttcagaagatggttgaaaatttaagaacagatttcagaattaaaaatcaaaaaccaacaatt
 ggatttggaaaatacagaacttagccaaaagaactctcaaaaccaggaaaaaactgcaagaacttaacacgtctaacagaaatgctatgccaga
 aggaaaaagagccaggaaacagtgattggaggaacgggaacaagagaagtttaactgaaagaagaactggaacgttgtaaaagtgcagtcct
 45 ccactttagtgtcttctcctggaggcggagctctctgaagttaaaatcacagaccatattgtgcaacaggaaaaccaccttctcaaatgaaatgga
 gaaaatgaaacagctgcacagatgtcccgtatctctgacttcagcaaaaaatctctagtgttctaagctacaacgaaaaactgctgaagaaaa
 ggaagctctgagtgaggaaatgaatagctgtgtc gataagttggcaaaaatcaagctcttttagagcatagaattgcgacgatgaagcaggaacagaa

atcctgggaacatcagagtgcgagcttaagtcacagctgggtggtctcagggaaagggtcagaatttagaagacaccgtgcagaatgtaaacc
 tgcaaatgtcccggatgaaatctgacctacgagtactcagcaggaaaaggaggctttaaacaagaagtgatgtctttacataagcaacttcaga
 atgctggtggcaagagctgggccccagagatagctactcatccatcagggctccataaccagcagaaaagggtgtcctgggacaagtggatca
 tctgatgaatgaggaacagcagctgctttggcaagagaatgagaggctccagaccatggtacagaacaccaaaggccgaactcacgcactccccg
 5 ggagaaggtccgtcaattggaatccaatcttctcccaagcaccaaaaacatctaaaccatcaggtaccatgaatcccacagagcaagaaaaatt
 gagcttaagagagagtgtatcagtttcagaaagaacaatctcctgctaacaggaaggtcagtcagatgaattcccttgaacaagaattagaac
 aattcatttggaaaaatgaaggcctgaaaaagaacaagtaaaactggatgagcagctcatggagatgcagcacctgaggtccactgcgacgcct
 agcccgtccctcatgcttgggatttgcagctgtccagcagcaagcctgtccgatggtgccaggaggagcagtttgcagctcaacgccagct
 gctgcaggcagaaaaggataaaccagcacctgcaggaggaacttgaaaacaggacctccgaaaccaacacaccacagggaaccaggaaca
 10 actgtaactgtcatggaggaacgaatgatagaagtgaacagaaactgaaactagtgaagggtcttcaagagaaagtgaatcagctcaag
 aacaactctgcaagaacactaaggcagacgcaatggtgaaggacttgtatgtgaaaatgccagttgttgaaagctctggaagtgactgaacag
 cgacagaaaacagcagagaagaaaaattacctcctggaggagaagattgccagcctcagtaatatagtttagaatctgacaccagcgccattga
 cttctacacctctttgaggtcatagccaaacaaagggtacactcatatttgcactttactgaaatagatgaacatttcagtaggttctcaacttaa
 aattaagcctaactaaaactgccagcaacacaactggagttccattatcataattagtttttaaatagaccttatgggagtttgaaaaataaac
 15 tcacataattcactacttaaatatttcccaagatttgaattatttttaaaatttaataagccaccaagaatgtggacatatgaaaattcaagaacctaaaaa
 ataccagtttgaatgagttttgtggttttggtttttaattattacaaatctatgtgtaaaatctagatattgaagtttgagatctgatgagaatggttgta
 taaactttattttaaaacaaatttaggtgttctacataattaaacttgaaagtcattataatagttttggttcttgaaattggtagacaatttagtagta
 taattggttaggagggcagggttattaagtggttattaaccgctgacatcagacaaacccaatctgtagaattctaacctcctaacacctgtgacag
 tattaccactcttctgtattatagatttagaactgatttactcaattgcactcttaactaatgttaaaagcttacttgctttaaacagccctttcttcttctt
 20 aaaagtttcatttggggagctggtcttctaagaaacggataaagccacataaattaaagcagttgaactagagggaagcactgaacaacacacttt
 ggagtaaatagctactcttagaaaaggaggaagcagaccatgtaggtttctgtcttcaaatcttagagttcataaatttacttgaggttgcctcaa
 gaactcagggaacaataactgtaaactgtcttctgaactactgtagggccttctaagaattgaaatgtataaaccatgtacctcatttattgtctta
 tataattacagccataactagaatttttattctacgttttagtaaatttaatttctgggggaaaaaaggccttgatttaggggttaaaaacctgacttatag
 aagagttatttaatataggtcaaaatttctgtgttcttattccttctataacctcaaatctgattctaagaatttcttactgtgataatcattggcatgccac
 25 ctgaggtcaaggagtgcacaataggactttccactcatgtcgaatcaaaactttatagaacagtcacatttttagattcggtaaccttttttcttcc
 aattataatctctgcttctagccacttccgccagcagttggtggaagacttactaggtgcagggcactttccaagttcatcacaacaacctgctgttt
 catgagacaataatccgaaaagttcgctttgatataattcctggagggccaagcccactatttacaanaagggtgaacagcaaaatcaagcactgcttt
 atgggcaggaacacaagagaaaagcaaaactgccaagaagtcacatgtcagaactcaatctcaacaaaataatttccatcagggaacttcagg
 gtttcttgggggcttatgagctcaccgggtcaaccagaggcctcactacaagagccttgacaaggcactgtttttgtgggactgggagttcaca
 30 ctgatgaagcaaaccttgaattttgcacagctctgtcagaagccctgagttcccttgataaagagtttaatttataccttccctataattatactt
 caaaatatttgacatctgctattatgccttcttagatcttcttctggtgcagacatttctagtaagtgttgactacttgtatggcattagctttcacag
 aaaattgttctacttaaaactgtggattggcctaggcctaaggacaaaaataaactaagtacctgtagtgtatttattgtatgtgtcaagttactcaa
 gtattgtgttggaactgaacaataatatttccagatagctggccttagcatgtgatcacggtgtgtatttttaattttgtctttacagtatgagag
 gtgtagggttaatttgttatttctataaattgtatttattgtatataaaatgtacaaatgaatgtaaataatgactttctgaaagtttagactacatttagaat
 35 ctctattcaaaatcaaaatgctgctcaaatgaatttaaccaacatctagggtgcttaatttctcattttatcccacttatgagattgggaaaaagatcaatat
 gagaataaccatacagataaccttaaatgtatgcatttgtgcaacaattttgagaaggtgagtggaatttataatttagttggcaatttataatagaact
 tatagcttttaaaagactttttaagacattaaatgtaaaacttaaaaatgttttagatctgtttcaactttacaatagcattctcaaaatattaagttatatat
 tttataggcatttagttgcttattaaaagcactgattttcaaaacttttgatttaagaacaattatttaagatcgtctcagaagatgggacttctgttcaaga
 aaagggaatcaagtttgcctttgagataatacgttacactaagaaaaggaaaatgtggatagtaaaacccacctctctcactcttactcttctt
 40 gctttttagaagcctgcacttaagcttagatttgaaggagagtagaaggaggagaagtagaaccacagtggtttatttttttctaaaactcttact
 aatccagatttttaaaactgttttaaatgtgaattctcccagaaattcaatgcattgcataatttagccttcggcatattttcatgaatagatcatgaagt
 cataggctccaaggcataggagagatctgcaggcttagtattttaaatatgcactattaccagggcagatattatgagaaactgtttcttcttaa
 ggggttatggcagacttgccttttaacatgtgagaaatgaattttttgtgatttattgtatttcttctgctgagtgaaaggagaaattgttgc
 tattgtcagcatcttaaggattttccagtcaaggcaaggctaagtgtttgtatagtattaagcaagtcattgtttgaatggattacctgtagtactc
 45 attggaatgatataattatacaagtaatgcaaaaaaccaagtcaaaagcctaattaaccaaagcactcatttaaaaatcatcatgtttggacctatctgg
 acctctcagcactgtaaaatagttttgttttggcatatgaatagctgtttaacaaatcaaagttagctttttgcttctcagctttttggcaatacaag
 ttaagttctaatggggagacattatcatggcatgacttaagggaacattggtttgtgaaggaaaaacagattatctaaagccatctctatgtttctgttc

agataaagattaatgagttctgtgtttatatacagctttgtatattcatcttagccattctatcctagaaagattttaatgtgagcttaagatgtaaataata
atfttgcaaacatgaaaaaaaaaaaaaaaaa (SEQ ID NO:3)

TABLE 12

Predicted Amino Acid Sequence of PN7771

5
10
15
20
25
30
35
40

MAEVTVPVRYVVFGIHCIMAKASSDVQVSGFHRKIQHVKNELCHMLSLEEVAPVLQQTLL
QDNLLGRVHFDQFKEALILILSRTLSEEHFQEPDCSLEAQPKYVRGGKRYGRRSLPEFQES
VEEFPEVTVIEPLDEEARPSHIPAGDCSEHWKTQRSEYEAEQQLRFWNPDDLNASQSGSSP
PQDWIEEKLQEVCELDGITRDGHLNRKKLVSICEQYGLQNVDGEMLEEVFHNLDPDGTMS
VEDFFYGLFKNGKSLTPSASTPYRQLKRHLMSQSFDESGRRTTTSSAMTSTIGFRVFSCLDD
GMGHASVERILDTWQEEGIENSQEILKALDFSLDGNINLTETLALENELLVTKNSIHQAAL
ASFKAIEIRHLLERVDQVVREKEKLRSDLDKAEKLSLMASEVDDHHAIEIRNEYNLRKL
DGEYKERIAALKNELRKEREQILQQAGKQRLELEQEIEKAKTEENYIRDRLALSLKENSRL
NELLENAEKLAEYENLTNKLQRNLENVLAEKFGDLDPSAEFFLQEERLTQMRNEYERQCR
VLQDQVDELQSELEEYRAQGRVLRPLKNSPSEEEVEANS GGIEPEHGLGSEECNPLNMSIEA
ELVIEQMKEQHHRDICCLRELEDKVRHYEKQLDET VVSCKKAQENMKQRHENETRLEK
QISDLKNEIAELQGQA AVLKEAHHEATCRHEEEKQLQVKLEEEKTHLQEKRLQHEMEL
KARLTQAQASFEREREGLOSSAWTEEKVRGLTQELEQFHQEQLTSLVEKHTLEKEELRKEL
LEKHQRELQEGREKMETECNRRTSQIEAQFQSDCQKVTERCESALQSLEGRYRQELKDLQE
QREEKSQWEFEKDELTQECAEAQELLKETLREKTTSLVLTQEREMLEKTYKEHLNSMV
VERQQLQDLEDLRNVSETQQSLSDQILELKSSHKRELREEREVL CQAGASEQLASQRLER
LEMEHDQERQEMMSKLLAMENIHKATCETADRERAEMSTEISRLQSKIKEMQQATSPLSM
LQSGCQVIGEEVEGDGALSLLQQGEQLLENGDVLLSLQRAHEQAVKENVKMATEISRLQ
QRLQKLEPGLVMSSCLDEPATEFFGNTAEQTEQFLQQNRTKQVEGVTRRHVLSLEDDEV
DLGSTGTSSVQRQEVKIEESEASVEGFSELENSEETRTESEWELKNQISQLQEQLMMLCADCD
RASEKKQDLLFDVSVLKKKLKMLERIEASPKYKLLYEDVSRENDCLQEELRMMETRYDE
ALENNKELTAEVFRLQDELKKMEEVTETFLSLEKSYDEVKIENEGLNVLVLRQKGKIEKLQE
SVVQRCDCCLWEASLENLEIEPDGNILQLNQTL EECVPRVRSVHHVIEECKQENQYLEGNT
QLLEKVKAEHIAWLHGTHQERPRVQNQVILEENTTL LGFQDKHFQHQATIAELELEKTK
LQELTRKLERVTILVKQKDVLSHGEEKKEELKAMMHDLQITCSEMQQKVELLRYESEKLQ
QENSILRNEITTLNEEDSISNLKLGTLNGSQEEMWQKTETVKQENAAVQKMVENLKKQISE
LKIKNQQLDLENTELS QKNSQNQEKLQELNQRLTEMLCQKEKEPGNSALEEREQEKFNLKE
ELERCKVQSSTLVSSLEAELSEVKIQTHIVQQENHLLKDELEKMKQLHRCPLDSDFQQKISS
VLSYNEKLLKEKEALSEELNSCVDKLAKSSLEHRIATMKQEQKSWEHQASLSQLVASQ
EKVQNLEDTVQNVNLQMSRMKSDLRVTQQEKEALKQEVMSLHKQLQNAGGKSWAPEIAT
HPSGLHNQQKRLSWDKLDHLMNEEQQLLWQENERLQTMVQNTKAELTHSREKVRQLESN
LLPKHQKHLNPSGTMNPTEQEKLSLKRECDQFQKEQSPANRKVSQMNSLEQELETIHLENE
GLKKKQVKLDEQLMEMQHRLSTATPSPSPAWDLQLLQQQACPMVPREQFLQLQRQLLQ
AERINQHLQEELNRTSETNTPQGNQEQLVTVM EERMIEVEQKLLVKRLLQEKVNQLKEQ
LCKNTKADAMVKDLYVENAQLLKALEVTEQRQKTA EKKNYLLEEKIASLSNIVRNLTAPL
TSTPPLRS (SEQ ID NO:4)

EXAMPLES 4-12

Identification of Protein-Protein Interactions

[0063] A yeast two-hybrid system as described in Example 1 using amino acids of the bait as set forth in Table 13 was performed. The clone that was identified by this procedure for each bait is set forth in Table 13 as the prey. The “AA” refers to the amino acids of the bait or prey. The
5 “NUC” refers to the nucleotides of the bait or prey. The Accession numbers refer to GB: GenBank accession numbers.

10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189
2190
2191
2192
2193
2194
2195
2196
2197
2198
2199
2200
2201
2202
2203
2204
2205
2206
2207
220

TABLE 13

Ex.	BAIT	ACCESSION	COORDINATES	PREY	ACCESSION	COORDINATES
4	LXR-alpha	GB: U22662	AA: 95-156	zyxin	GB: X94991	AA 323-572
5	LXR-alpha	GB: U22662	AA 95-156	LIMS1	GB: U09284	AA 25-90
6	LXR-alpha	GB: U22662	AA 95-277	Homer-3	GB: NM_004838	AA 182-354
7	LXR-alpha	GB: U22662	AA 95-277	RACK1	GB: M24194	AA 178-317
8	LXR-alpha	GB: U22662	AA 257-448	EIF3S1	GB: NM_003758	AA 94-254
9	LXR-alpha	GB: U22662	AA 156-448	EIF3S1	GB: NM_003758	AA 94-254
10	LXR-alpha	GB: U22662	AA 95-277	PSMD11	GB: NM_002815	AA 87-422
11	LXR-alpha	GB: U22662	AA 257-448	KIAA0610	GB: AB011182	AA 36-245
12	LXR-alpha	GB: U22662	AA 156-448	CIR	GB: U03644	AA 226-450

EXAMPLE 13

Generation of Polyclonal Antibody Against Protein Complexes

[0064] As shown above, LXR-alpha interacts with utrophin to form a complex. A complex of the two proteins is prepared, e.g., by mixing purified preparations of each of the two proteins.

5 If desired, the protein complex can be stabilized by cross-linking the proteins in the complex, by methods known to those of skill in the art. The protein complex is used to immunize rabbits and mice using a procedure similar to that described by Harlow et al. (1988). This procedure has been shown to generate Abs against various other proteins (for example, see Kraemer et al., 1993).

[0065] Briefly, purified protein complex is used as immunogen in rabbits. Rabbits are
10 immunized with 100 µg of the protein in complete Freund's adjuvant and boosted twice in three-week intervals, first with 100 µg of immunogen in incomplete Freund's adjuvant, and followed by 100 µg of immunogen in PBS. Antibody-containing serum is collected two weeks thereafter. The antisera is preadsorbed with LXR-alpha and utrophin, such that the remaining antisera comprises antibodies which bind conformational epitopes, i.e., complex-specific epitopes, present on the LXR-alpha-utrophin complex but not on the monomers.

[0066] Polyclonal antibodies against each of the complexes set forth in Tables 1-10 are prepared in a similar manner by mixing the specified proteins together, immunizing an animal and isolating antibodies specific for the protein complex, but not for the individual proteins.

[0067] Polyclonal antibodies against the protein set forth in Table 12 are prepared in a
20 similar manner by immunizing an animal with the protein and isolating antibodies specific for the protein.

EXAMPLE 14

Generation of Monoclonal Antibodies Specific for Protein Complexes

25 [0068] Monoclonal antibodies are generated according to the following protocol. Mice are immunized with immunogen comprising LXR-alpha/utrophin complexes conjugated to keyhole limpet hemocyanin using glutaraldehyde or EDC as is well known in the art. The complexes can be prepared as described in Example 13, and may also be stabilized by cross-linking. The immunogen is mixed with an adjuvant. Each mouse receives four injections of 10 to 100 µg of
30 immunogen, and after the fourth injection blood samples are taken from the mice to determine if the serum contains antibody to the immunogen. Serum titer is determined by ELISA or RIA. Mice

with sera indicating the presence of antibody to the immunogen are selected for hybridoma production.

[0069] Spleens are removed from immune mice and a single-cell suspension is prepared (Harlow et al., 1988). Cell fusions are performed essentially as described by Kohler et al. (1975). Briefly, P3.65.3 myeloma cells (American Type Culture Collection, Rockville, MD) or NS-1 myeloma cells are fused with immune spleen cells using polyethylene glycol as described by Harlow et al. (1988). Cells are plated at a density of 2×10^5 cells/well in 96-well tissue culture plates. Individual wells are examined for growth, and the supernatants of wells with growth are tested for the presence of LXR-alpha/utrophin complex-specific antibodies by ELISA or RIA using LXR-alpha/utrophin complex as target protein. Cells in positive wells are expanded and subcloned to establish and confirm monoclonality.

[0070] Clones with the desired specificities are expanded and grown as ascites in mice or in a hollow fiber system to produce sufficient quantities of antibodies for characterization and assay development. Antibodies are tested for binding to LXR-alpha alone or to utrophin alone, to determine which are specific for the LXR-alpha/utrophin complex as opposed to those that bind to the individual proteins.

[0071] Monoclonal antibodies against each of the complexes set forth in Tables 1-10 are prepared in a similar manner by mixing the specified proteins together, immunizing an animal, fusing spleen cells with myeloma cells and isolating clones which produce antibodies specific for the protein complex, but not for the individual proteins.

[0072] Monoclonal antibodies against the protein set forth in Table 12 are prepared in a similar manner by immunizing an animal with the protein, fusing spleen cells with myeloma cells and isolating clones which produce antibodies specific for the protein.

EXAMPLE 15

In vitro Identification of Modulators for Protein-Protein Interactions

[0073] The present invention is useful in screening for agents that modulate the interaction of LXR-alpha and utrophin. The knowledge that LXR-alpha and utrophin form a complex is useful in designing such assays. Candidate agents are screened by mixing LXR-alpha and utrophin (a) in the presence of a candidate agent, and (b) in the absence of the candidate agent. The amount of complex formed is measured for each sample. An agent modulates the interaction of LXR-alpha and utrophin if the amount of complex formed in the presence of the agent is greater than

(promoting the interaction), or less than (inhibiting the interaction) the amount of complex formed in the absence of the agent. The amount of complex is measured by a binding assay, which shows the formation of the complex, or by using antibodies immunoreactive to the complex.

[0074] Briefly, a binding assay is performed in which immobilized LXR-alpha is used to bind labeled utrophin. The labeled utrophin is contacted with the immobilized LXR-alpha under aqueous conditions that permit specific binding of the two proteins to form a LXR-alpha/utrophin complex in the absence of an added test agent. Particular aqueous conditions may be selected according to conventional methods. Any reaction condition can be used as long as specific binding of LXR-alpha/utrophin occurs in the control reaction. A parallel binding assay is performed in which the test agent is added to the reaction mixture. The amount of labeled utrophin bound to the immobilized LXR-alpha is determined for the reactions in the absence or presence of the test agent.

If the amount of bound, labeled utrophin in the presence of the test agent is different than the amount of bound labeled utrophin in the absence of the test agent, the test agent is a modulator of the interaction of LXR-alpha and utrophin.

[0075] Candidate agents for modulating the interaction of each of the protein complexes set forth in Tables 1-10 are screened *in vitro* in a similar manner.

EXAMPLE 16

In vivo Identification of Modulators for Protein-Protein Interactions

[0076] In addition to the *in vitro* method described in Example 15, an *in vivo* assay can also be used to screen for agents which modulate the interaction of LXR-alpha and utrophin. Briefly, a yeast two-hybrid system is used in which the yeast cells express (1) a first fusion protein comprising LXR-alpha or a fragment thereof and a first transcriptional regulatory protein sequence, e.g., GAL4 activation domain, (2) a second fusion protein comprising utrophin or a fragment thereof and a second transcriptional regulatory protein sequence, e.g., GAL4 DNA-binding domain, and (3) a reporter gene, e.g., β -galactosidase, which is transcribed when an intermolecular complex comprising the first fusion protein and the second fusion protein is formed. Parallel reactions are performed in the absence of a test agent as the control and in the presence of the test agent. A functional LXR-alpha/utrophin complex is detected by detecting the amount of reporter gene expressed. If the amount of reporter gene expression in the presence of the test agent is different than the amount of reporter gene expression in the absence of the test agent, the test agent is a modulator of the interaction of LXR-alpha and utrophin.

[0077] Candidate agents for modulating the interaction of each of the protein complexes set forth in Tables 1-10 are screened *in vivo* in a similar manner.

[0078] While the invention has been disclosed in this patent application by reference to the details of preferred embodiments of the invention, it is to be understood that the disclosure is intended in an illustrative rather than in a limiting sense, as it is contemplated that modifications will readily occur to those skilled in the art, within the spirit of the invention and the scope of the appended claims.

BIBLIOGRAPHY

- Aronheim et al., 1997. Isolation of an AP-1 repressor by a novel method for detecting protein-protein interactions. *Mol. Cell. Biol.* **17**:3094-3102.
- Altschul, S.F. et al. (1990). Basic local alignment search tool. *J. Mol. Biol.* **215**:403-410.
- Altschul, S.F. et al. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucl. Acids Res.* **25**:3389-3402.
- Banres, P.J. (1998). Anti-inflammatory actions of glucocorticoids: molecular mechanisms. *Clinical Science* **94**:557-572.
- Bartel, P.L. et al. (1993). "Using the 2-hybrid system to detect protein-protein interactions." In: *Cellular Interactions in Development: A Practical Approach*, Oxford University Press, pp. 153-179.
- Bartel, P.L. et al. (1996). A protein linkage map of Escherichia coli bacteriophage T7. *Nat Genet* **12**:72-77.
- Bartel, P.L. and Fields, S. (1997). *The Yeast Two-Hybrid System*. New York: Oxford University Press.
- Brown, A.J. and Jessup, W. (1999). Oxysterols and atherosclerosis. *Atherosclerosis* **142**:1-28.
- Chevray, P.M. and Nathans, D.N. (1992). Protein interaction cloning in yeast: identification of mammalian proteins that interact with the leucine zipper of Jun. *Proc. Natl. Acad. Sci. USA* **89**:5789-5793.
- DeFranco, D. et al. (1998). Molecular chaperones and subcellular trafficking of steroid receptors. *J. Steroid Biochem. and Mol. Biol.* **65**:51-58.
- Duprez, D. et al. (2000). Aldosterone and vascular damage. *Curr. Hypertens. Rep.* **2**:327-34.
- Fields S and Song O-K (1989). A novel genetic system to detect protein-protein interactions. *Nature* **340**:245-246.
- Geller, D.S. et al. (2000). Activating mineralocorticoid receptor mutation in hypertension exacerbated by pregnancy. *Science* **289**:119-123.
- Harlow et al. (1988). *Antibodies: A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.).

- Haynes, M.P. et al. (2000). Molecular mechanisms of estrogen actions on the vasculature. *J. Nucl. Cardiol.* **7**:500-508.
- Hong, Y.R. et al. (2000). Cloning and characterization of a novel human ninein protein that interacts with the glycogen synthase kinase 3 β . *Biochim Biophys Acta* **1492**:513-6.
- 5 Hsieh, J.J. et al. (1999). CIR, a corepressor linking the DNA binding factor CBF1 to the histone deacetylase complex. *Proc Natl Acad Sci USA.* **96**:23-8.
- Klein-Hitpass, L. et al. (1998). Targets of activated steroid hormone receptors: basal transcription factors and receptor interacting proteins. *J. Mol. Med.* **76**:490-496.
- 10 Kohler, G. and Milstein, C. (1975). Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature* **256**:495-497.
- Kraemer, F.B. et al. (1993). Detection of hormone-sensitive lipase in various tissues. I. Expression of an HSL/bacterial fusion protein and generation of anti-HSL antibodies. *J. Lipid Res.* **34**:663-672.
- 15 Leanna, C.A. and Hannink, M. (1996). The reverse two-hybrid system: a genetic scheme for selection against specific protein/protein interactions. *Nucl. Acids Res.* **24**:3341-3347.
- Mack, G.J. et al. (1998). Autoantibodies to a group of centrosomal proteins in human autoimmune sera reactive with the centrosome. *Arthritis Rheum.* **41**:551-8.
- Mangelsdorf, D. et al. (1995). The nuclear receptor superfamily: the second decade. *Cell* **83**:835-839.
- 20 Moras, D. and Gronemeyer, H. (1998). The nuclear receptor ligand-binding domain: structure and function. *Curr. Op. Cell Biol.* **10**:384-391.
- Schroepfer, G.J. (2000). Oxysterols: modulators of cholesterol metabolism and other processes. *Physiol. Rev.* **80**:361-554.
- 25 Tu, Y. et al. (1998). Nck-2, a novel Src homology2/3-containing adaptor protein that interacts with the LIM-only protein PINCH and components of growth factor receptor kinase-signaling pathways. *Mol Biol Cell.* **9**:3367-82.
- Wetmur, J.G. and Davidson, N. (1968). "Kinetics of renaturation of DNA." *J. Mol. Biol.* **31**:349-370.
- 30 Whitfield, G. et al. (1999). Steroid hormone receptors: evolution, ligands, and molecular basis of biologic function. *J. Biol. Chem. Suppl* **32-33**:110-122.
- Wolffe, A. et al. (1997). Chromatin remodeling by steroid and nuclear receptors. *Cell Research* **7**:127-142.

PCT Published Application No. WO 97/27296

35 PCT Published Application No. WO 99/65939

U.S. Patent No. 5,622,852

U.S. Patent No. 5,773,218